Combining Covariate Adjustment with Information Adaptive Designs

UNIVERSITEIT GENT

Kelly Van Lancker

Joint work with

Michael Rosenblum and Josh Betz

JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

Planning

1. Background, Problem Setting and Set Up
2. Potential Solution: Information-Adaptive Design
3. Simulation Study

Outline

Covariate Adjustment

- **Covariate adjustment** is a statistical analysis method with high potential to **improve precision** for many trials.
  - **Pre-planned** adjustment for baseline variables when estimating **average treatment effect**.
  - Estimand is same as when using unadjusted estimator (e.g., difference in means).
  - **Goal**: avoid making any model assumptions beyond what's assumed for unadjusted estimator (**robustness to model misspecification**).

(e.g., Koch et al., 1998; Yang and Tsiatis, 2001; Rubin and van der Laan, 2008; Tsiatis et al., 2008; Moore and van der Laan, 2009b,a; Zhang, 2015; Jiang et al., 2018; Benkeser et al., 2020)

FDA Guidance: Example

- Primary endpoint Y: binary.

FDA Guidance: Example

- Primary endpoint Y: binary.
- Estimand: $\theta = E(Y|A=1) - E(Y|A=0)$.

FDA Guidance: Example

- Primary endpoint Y: binary.
- Estimand: $\theta = E(Y|A=1) - E(Y|A=0)$.
- Estimator: G-computation/Standardization
  1. Fit logistic regression model for

$$P(Y=1|A,B) = logit^{-1}(\gamma_0 + \gamma_1 A + \gamma_2 B).$$

  2. Compute standardized estimators for treatment specific means
     - $\hat{E}(Y|A=1) = \frac{1}{n}\sum_{i=1}^{n} logit^{-1}(\hat{\gamma}_0 + \hat{\gamma}_1 + \hat{\gamma}_2 B_i)$
     - $\hat{E}(Y|A=0) = \frac{1}{n}\sum_{i=1}^{n} logit^{-1}(\hat{\gamma}_0 + \hat{\gamma}_2 B_i)$
  3. Calculate $\hat{\theta} = \hat{E}(Y|A=1) - \hat{E}(Y|A=0)$

Potential Challenge

- The **uncertainty** at the design stage about the **amount of precision gain and corresponding sample size reduction**.

Potential Challenge

- The **uncertainty** at the design stage about the **amount of precision gain and corresponding sample size reduction**.
  - **Approach 1**: assume conservatively that covariate adjustment will not lead to a precision gain.

Potential Challenge

- The **uncertainty** at the design stage about the **amount of precision gain and corresponding sample size reduction**.
  - **Approach 1**: assume conservatively that covariate adjustment will not lead to a precision gain.
  - **Approach 2**: consider how much precision can be gained based on external (trial) data when calculating the sample size. (Li et al., 2023)
    - An incorrect projection of a covariate's prognostic value, may still lead to an over- or underpowered future trial.

Potential Challenge

- The **uncertainty** at the design stage about the **amount of precision gain and corresponding sample size reduction**.
  - **Approach 1**: assume conservatively that covariate adjustment will not lead to a precision gain.
  - **Approach 2**: consider how much precision can be gained based on external (trial) data when calculating the sample size. (Li et al., 2023)
    - An incorrect projection of a covariate's prognostic value, may still lead to an over- or underpowered future trial.
  - **Potential solution**: combine covariate adjustment with **information-adaptive designs** (also known as information monitoring).

Outline

Algorithm for Analysis Timing: Design Stage

- Specify the operating characteristics of the study

Algorithm for Analysis Timing: Design Stage

- Specify the operating characteristics of the study
- We **compute the maximum/total information** needed to preserve these operational characteristics

$$\left(\frac{z_{\alpha/2} + z_{\beta}}{\theta_A - \theta_0}\right)^2,$$

for a fixed design (no interim analyses), and

$$\left(\frac{z_{\alpha/2} + z_{\beta}}{\theta_A - \theta_0}\right)^2 IF$$

when data is sequentially monitored with the possibility of early stopping.

(Mehta and Tsiatis, 2001)

Algorithm for Analysis Timing: Information

- We propose to **monitor the accrued information**, $(\widehat{se}(\hat{\theta}_t))^{-2}$, through time t.

Algorithm for Analysis Timing: Information

- We propose to **monitor the accrued information**, $(\widehat{se}(\hat{\theta}_t))^{-2}$, through time t.
- We consider a trial with an interim analysis when 50% of the information is available:

Algorithm for Analysis Timing: Information

- We propose to **monitor the accrued information**, $(\widehat{se}(\hat{\theta}_t))^{-2}$, through time t.
- We consider a trial with an interim analysis when 50% of the information is available:
  - We conduct the interim analysis at time t_1 when

$$(\widehat{se}(\hat{\theta}_{t_1}))^{-2} \geq 0.5 \cdot \left(\frac{z_{\alpha/2} + z_{\beta}}{\theta_A - \theta_0} \right)^2 IF.$$

Algorithm for Analysis Timing: Information

- We propose to **monitor the accrued information**, $(\widehat{se}(\hat{\theta}_t))^{-2}$, through time t.
- We consider a trial with an interim analysis when 50% of the information is available:
  - We conduct the interim analysis at time t_1 when

$$(\widehat{se}(\hat{\theta}_{t_1}))^{-2} \geq 0.5 \cdot \left(\frac{z_{\alpha/2} + z_{\beta}}{\theta_A - \theta_0} \right)^2 IF.$$

  - We conduct the final analysis at time t_2 when

$$(\widehat{se}(\hat{\theta}_{t_2}))^{-2} \geq \left(\frac{z_{\alpha/2} + z_{\beta}}{\theta_A - \theta_0} \right)^2 IF.$$

(Mehta and Tsiatis, 2001; Zhang, 2009)

Algorithm for Analysis Timing: (Dis)advantages

- The **information-adaptive design** is well suited for being adopted for covariate adjusted estimators:
  - We do **not have to prespecify the prognostic value of the covariates** nor other nuisance parameters.
  - When the estimator is more efficient than unadjusted estimator, covariate adjustment can lead to a **shorter trial** due to faster information accrual.

Algorithm for Analysis Timing: (Dis)advantages

- The **information-adaptive design** is well suited for being adopted for covariate adjusted estimators:
  - We do **not have to prespecify the prognostic value of the covariates** nor other nuisance parameters.
  - When the estimator is more efficient than unadjusted estimator, covariate adjustment can lead to a **shorter trial** due to faster information accrual.
- **Administrative inconvenience**: it does not give an idea to the investigators about the necessary resources (i.e., length of study, sample size, ...).

Algorithm for Analysis Timing: Practical Issues

- We suggest to posit some **guesses on the nuisance parameters**.
  - Probability of success in control arm (binary endpoint), prognostic value of covariates, ...

Algorithm for Analysis Timing: Practical Issues

- We suggest to posit some **guesses on the nuisance parameters**.
  - Probability of success in control arm (binary endpoint), prognostic value of covariates, ...
- Assessing feasibility by **estimating the number of participants** corresponding with the maximum information.
  - Usings the standard formulas for sample size calculations.
  - We recommend setting the **sample size conservatively** as if there were no precision gain from covariate adjustment.

Algorithm for Analysis Timing: Practical Issues

- We suggest to posit some **guesses on the nuisance parameters**.
  - Probability of success in control arm (binary endpoint), prognostic value of covariates, ...
- Assessing feasibility by **estimating the number of participants** corresponding with the maximum information.
  - Usings the standard formulas for sample size calculations.
  - We recommend setting the **sample size conservatively** as if there were no precision gain from covariate adjustment.
- However, **miscalculations can occur** at the design stage.
  - We should use the **emerging data to evaluate** whether the maximum information will be reached in the planned time.

Algorithm for Analysis Timing: Practical Issues

- However, **miscalculations can occur** at the design stage.

Algorithm for Analysis Timing: Practical Issues

- However, **miscalculations can occur** at the design stage.
- We should use the **emerging data to evaluate** whether the maximum information will be reached with the planned sample size.

Algorithm for Analysis Timing: Practical Issues

- However, **miscalculations can occur** at the design stage.
- We should use the **emerging data to evaluate** whether the maximum information will be reached with the planned sample size.
- If not, we should **update the maximum sample size** at time t as

$$n_{max} = n(t)\frac{\left(\frac{z_{\alpha/2}+z_{\beta}}{\theta_A-\theta_0}\right)^2 IF}{(\widehat{se}(\hat{\theta}_t))^{-2}},$$

where $n(t)$ is the number of patients used in the analysis at time t.

(Mehta and Tsiatis, 2001)

Algorithm for Analysis Timing: Contribution

- We proposed an **information-adaptive trial design**, where the analysis timing is based on **accruing information** and is **data-adaptive**.

Algorithm for Analysis Timing: Contribution

- We proposed an **information-adaptive trial design**, where the analysis timing is based on **accruing information** and is **data-adaptive**.
- By **automatically adapting to amount of precision gain** due to covariate adjustment, it results in **correctly powered** trials.

Algorithm for Analysis Timing: Contribution

- We proposed an **information-adaptive trial design**, where the analysis timing is based on **accruing information** and is **data-adaptive**.
- By **automatically adapting to amount of precision gain** due to covariate adjustment, it results in **correctly powered** trials.
- Information will accrue faster as covariate adjusted estimators typically have smaller variance, leading to **faster trials at no additional cost**.

Algorithm for Analysis Timing: Contribution

- We proposed an **information-adaptive trial design**, where the analysis timing is based on **accruing information** and is **data-adaptive**.
- By **automatically adapting to amount of precision gain** due to covariate adjustment, it results in **correctly powered** trials.
- Information will accrue faster as covariate adjusted estimators typically have smaller variance, leading to **faster trials at no additional cost**.
- Since adaptations to the analysis timing are **pre-planned** based on nuisance parameters only, they are generally **acceptable to regulators**.

Outline

MISTIE III trial (Stroke)

- Functional outcome: proportion of patients who achieved a modified Rankin Scale score of 0-3 at 365 days (**binary**).
- Estimand of interest: **risk difference**.
- Total sample size of approximately 498 patients (in original trial):

MISTIE III trial (Stroke)

- Functional outcome: proportion of patients who achieved a modified Rankin Scale score of 0-3 at 365 days (**binary**).
- Estimand of interest: **risk difference**.
- Total sample size of approximately 498 patients (in original trial):
  - 1:1 randomization
  - Power of 88% to detect an average effect size of 13% at a 5% significance level
  - Success rate: 25% in standard medical care group versus 38% in MISTIE group

MISTIE III trial (Stroke)

- Functional outcome: proportion of patients who achieved a modified Rankin Scale score of 0-3 at 365 days (**binary**).
- Estimand of interest: **risk difference**.
- Total sample size of approximately 498 patients (in original trial):
  - 1:1 randomization
  - Power of 88% to detect an average effect size of 13% at a 5% significance level
  - Success rate: 25% in standard medical care group versus 38% in MISTIE group
- We will focus on information instead of sample size!

Simulation Study: $K = 1$

- **Information-adaptive design** with maximum information equal to 582
- **Maximum sample size design** with $n_{max} = 498$

		$\theta = 0.13$ (Alternative)			
		Power	**ASN**	AAT	AI
Information-adaptive design	Unadjusted	88.4%	**571**	1876	582
	Standardization	87.3%	**433**	1509	567
Maximum sample size design	Unadjusted	83.1%	-	1682	508
	Standardization	91.1%	-	1682	652

ASN: average sample number; AAT: average analysis time (days); AI: average information.

Conclusion under alternative:

24% reduction of sample size due to covariate adjustment

Simulation Study

- **Information-adaptive design** with maximum information equal to 582
- **Maximum sample size design** with $n_{max} = 498$

		$\theta = 0$ (Null)			
		Type I	**ASN**	AAT	AI
Information-adaptive design	Unadjusted	5.28%	**569**	1871	582
	Standardization	5.28%	**402**	1427	568
Maximum sample size design	Unadjusted	5.14%	-	1682	509
	Standardization	5.14%	-	1682	705

ASN: average sample number; AAT: average analysis time (days); AI: average information.

Conclusion under null:

29% reduction of sample size due to covariate adjustment

Thank you for your attention!

Interested? https://doi.org/10.48550/arXiv.2201.12921
E-mail: kelly.vanlancker@ugent.be
Website: kellyvanlancker.com

References I

Benkeser, D., I. Díaz, A. Luedtke, J. Segal, D. Scharfstein, and M. Rosenblum (2020). Improving precision and power in randomized trials for covid-19 treatments using covariate adjustment, for binary, ordinal, and time-to-event outcomes. *Biometrics*.

Jiang, F., L. Tian, H. Fu, T. Hasegawa, and L. J. Wei (2018). Robust alternatives to ANCOVA for estimating the treatment effect via a randomized comparative study. *Journal of the American Statistical Association 0*, 1–37.

Koch, G. G., C. M. Tangen, J.-W. Jung, and I. A. Amara (1998). Issues for covariance analysis of dichotomous and ordered categorical data from randomized clinical trials and non-parametric strategies for addressing them. *Stat. Med. 17*(15-16), 1863–1892.

References II

Li, X., S. Li, and A. Luedtke (2023). Estimating the efficiency gain of covariate-adjusted analyses in future clinical trials using external data. *Journal of the Royal Statistical Society Series B: Statistical Methodology 85*(2), 356–377.

Mehta, C. R. and A. A. Tsiatis (2001). Flexible sample size considerations using information-based interim monitoring. *Drug information journal: DIJ/Drug Information Association 35*(4), 1095–1112.

Moore, K. and M. J. van der Laan (2009a). Covariate adjustment in randomized trials with binary outcomes: Targeted maximum likelihood estimation. *Stat. Med. 28*(1), 39–64.

Moore, K. L. and M. J. van der Laan (2009b). Increasing power in randomized trials with right censored outcomes through covariate adjustment. *Journal of Biopharmaceutical Statistics 19*(6), 1099–1131. PMID: 20183467.

References III

Rubin, D. and M. van der Laan (2008). Covariate adjustment for the intention-to-treat parameter with empirical efficiency maximization. *U.C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 229*, https://biostats.bepress.com/ucbbiostat/paper229.

Tsiatis, A. A., M. Davidian, M. Zhang, and X. Lu (2008). Covariate adjustment for two-sample treatment comparisons in randomized clinical trials: a principled yet flexible approach. *Statistics in medicine 27*(23), 4658–4677.

Yang, L. and A. Tsiatis (2001). Efficiency study of estimators for a treatment effect in a pretest-posttest trial. *The American Statistician 55*(4), 314–321.

Zhang, D. (2009). Lecture notes for statistical principles of clinical trials (modified from dr. a. tsiatis' lecture notes).

References IV

Zhang, M. (2015, Jan). Robust methods to improve efficiency and reduce bias in estimating survival curves in randomized clinical trials. *Lifetime Data Analysis 21*(1), 119–137.